

RAW SEQUENCE LISTING
PATENT APPLICATION US/09/945,1820590
1205
OIPEDATE: 01/11/2002
TIME: 22:33:09

INPUT SET: S36718.raw

This Raw Listing contains the General
Information Section and up to the first 5 pages.

SEQUENCE LISTING

2.

ENTERED

1
2
3 (1) General Information:
4
5 (i) APPLICANT: Celeste, Anthony J.
6 Wozney, John
7 Rosen, Vicki A.
8 Wolfman, Neil
9 Thomsen, Gerald H.
10 Melton, Douglas A.
11
12 (ii) TITLE OF INVENTION: TENDON-INDUCING COMPOSITIONS
13
14 (iii) NUMBER OF SEQUENCES: 35
15
16 (iv) CORRESPONDENCE ADDRESS:
17 (A) ADDRESSEE: GENETICS INSTITUTE, INC.
18 (B) STREET: 87 CambridgePark Drive
19 (C) CITY: Cambridge
20 (D) STATE: Massachusetts
21 (E) COUNTRY: USA
22 (F) ZIP: 02140
23
24 (v) COMPUTER READABLE FORM:
25 (A) MEDIUM TYPE: Floppy disk
26 (B) COMPUTER: IBM PC compatible
27 (C) OPERATING SYSTEM: PC-DOS/MS-DOS
28 (D) SOFTWARE: PatentIn Release #1.0, Version #1.25
29
30 (vi) CURRENT APPLICATION DATA:
31 (A) APPLICATION NUMBER: 09/945,182
32 (B) FILING DATE:
33 (C) CLASSIFICATION:
34
35 (vii) PRIOR APPLICATION DATA:
36 (A) APPLICATION NUMBER: 08/808,324
37 (B) FILING DATE:
38
39
40 (viii) ATTORNEY/AGENT INFORMATION:
41 (A) NAME: Lazar, Steven R.
42 (B) REGISTRATION NUMBER: 32,618
43 (C) REFERENCE/DOCKET NUMBER: 5202-D
44
45 (ix) TELECOMMUNICATION INFORMATION:
46 (A) TELEPHONE: 617 498-8260

RAW SEQUENCE LISTING PATENT APPLICATION US/09/945,182

DATE: 01/11/2002
TIME: 22:33:09

INPUT SET: S36718.raw

```

47         (B) TELEFAX: 617 876-5851
48
49
50 (2) INFORMATION FOR SEQ ID NO:1:
51
52     (i) SEQUENCE CHARACTERISTICS:
53         (A) LENGTH: 926 base pairs
54         (B) TYPE: nucleic acid
55         (C) STRANDEDNESS: single
56         (D) TOPOLOGY: linear
57
58     (ii) MOLECULE TYPE: DNA (genomic)
59
60     (vi) ORIGINAL SOURCE:
61         (A) ORGANISM: Homo sapiens
62
63     (vii) IMMEDIATE SOURCE:
64         (B) CLONE: v1-1
65
66     (ix) FEATURE:
67         (A) NAME/KEY: mat_peptide
68         (B) LOCATION: 571..882
69
70     (ix) FEATURE:
71         (A) NAME/KEY: CDS
72         (B) LOCATION: 1..882
73
74
75     (xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:
76
77 GCG CGT AAT ACG ACT CAC TAT AGG GCG AAT TGG GTA CGG GGC CCA GGC      48
78 Ala Arg Asn Thr Thr His Tyr Arg Ala Asn Trp Val Arg Gly Pro Gly
79 -190                      -185                      -180                      -175
80
81 AGC TGG ACT TCT CCG CCG TTG CTG CTG CTG TCC ACG TGC CCG GGC GCC      96
82 Ser Trp Thr Ser Pro Pro Leu Leu Leu Ser Thr Cys Pro Gly Ala
83                      -170                      -165                      -160
84
85 GCC CGA GCG CCA CGC CTG CTG TAC TCG CGG GCA GCT GAG CCC CTA GTC      144
86 Ala Arg Ala Pro Arg Leu Leu Tyr Ser Arg Ala Ala Glu Pro Leu Val
87                      -155                      -150                      -145
88
89 GGT CAG CGC TGG GAG GCG TTC GAC GTG GCG GAC GCC ATG AGG CGC CAC      192
90 Gly Gln Arg Trp Glu Ala Phe Asp Val Ala Asp Ala Met Arg Arg His
91                      -140                      -135                      -130
92
93 CGT CGT GAA CCG CGC CCC CCC CGC GCG TTC TGC CTC TTG CTG CGC GCA      240
94 Arg Arg Glu Pro Arg Pro Pro Arg Ala Phe Cys Leu Leu Leu Arg Ala
95                      -125                      -120                      -115
96
97 GTG GCA GGC CCG GTG CCG AGC CCG TTG GCA CTG CGG CGA CTG GGC TTC      288
98 Val Ala Gly Pro Val Pro Ser Pro Leu Ala Leu Arg Arg Leu Gly Phe
99 -110                      -105                      -100                      -95

```

INPUT SET: S36718.raw

100																		
101	GGC	TGG	CCG	GGC	GGA	GGG	GGC	TCT	GCG	GCA	GAG	GAG	CGC	GCG	GTG	CTA		336
102	Gly	Trp	Pro	Gly	Gly	Gly	Gly	Ser	Ala	Ala	Glu	Glu	Arg	Ala	Val	Leu		
103					-90					-85					-80			
104																		
105	GTC	GTC	TCC	TCC	CGC	ACG	CAG	AGG	AAA	GAG	AGC	TTA	TTC	CGG	GAG	ATC		384
106	Val	Val	Ser	Ser	Arg	Thr	Gln	Arg	Lys	Glu	Ser	Leu	Phe	Arg	Glu	Ile		
107				-75					-70					-65				
108																		
109	CGC	GCC	CAG	GCC	CGC	GCG	CTC	GGG	GCC	GCT	CTG	GCC	TCA	GAG	CCG	CTG		432
110	Arg	Ala	Gln	Ala	Arg	Ala	Leu	Gly	Ala	Ala	Leu	Ala	Ser	Glu	Pro	Leu		
111			-60					-55					-50					
112																		
113	CCC	GAC	CCA	GGA	ACC	GGC	ACC	GCG	TCG	CCA	AGG	GCA	GTC	ATT	GGC	GGC		480
114	Pro	Asp	Pro	Gly	Thr	Gly	Thr	Ala	Ser	Pro	Arg	Ala	Val	Ile	Gly	Gly		
115		-45					-40					-35						
116																		
117	CGC	AGA	CGG	AGG	AGG	ACG	GCG	TTG	GCC	GGG	ACG	CGG	ACA	GCG	CAG	GGC		528
118	Arg	Arg	Arg	Arg	Arg	Thr	Ala	Leu	Ala	Gly	Thr	Arg	Thr	Ala	Gln	Gly		
119	-30					-25					-20					-15		
120																		
121	AGC	GGC	GGG	GGC	GCG	GGC	CGG	GGC	CAC	GGG	CGC	AGG	GGC	CGG	AGC	CGC		576
122	Ser	Gly	Gly	Gly	Ala	Gly	Arg	Gly	His	Gly	Arg	Arg	Gly	Arg	Ser	Arg		
123				-10					-5						1			
124																		
125	TGC	AGC	CGC	AAG	CCG	TTG	CAC	GTG	GAC	TTC	AAG	GAG	CTC	GGC	TGG	GAC		624
126	Cys	Ser	Arg	Lys	Pro	Leu	His	Val	Asp	Phe	Lys	Glu	Leu	Gly	Trp	Asp		
127			5					10					15					
128																		
129	GAC	TGG	ATC	ATC	GCG	CCG	CTG	GAC	TAC	GAG	GCG	TAC	CAC	TGC	GAG	GGC		672
130	Asp	Trp	Ile	Ile	Ala	Pro	Leu	Asp	Tyr	Glu	Ala	Tyr	His	Cys	Glu	Gly		
131		20					25					30						
132																		
133	CTT	TGC	GAC	TTC	CCT	TTG	CGT	TCG	CAC	CTC	GAG	CCC	ACC	AAC	CAT	GCC		720
134	Leu	Cys	Asp	Phe	Pro	Leu	Arg	Ser	His	Leu	Glu	Pro	Thr	Asn	His	Ala		
135	35					40					45					50		
136																		
137	ATC	ATT	CAG	ACG	CTG	CTC	AAC	TCC	ATG	GCA	CCA	GAC	GCG	GCG	CCG	GCC		768
138																		

RAW SEQUENCE LISTING
PATENT APPLICATION US/09/945,182DATE: 01/11/2002
TIME: 22:33:10

INPUT SET: S36718.raw

926

153 GCGGCCGAGG ATCC

154

155

156 (2) INFORMATION FOR SEQ ID NO:2:

157

158 (i) SEQUENCE CHARACTERISTICS:

159 (A) LENGTH: 294 amino acids

160 (B) TYPE: amino acid

161 (D) TOPOLOGY: linear

162

163 (ii) MOLECULE TYPE: protein

164

165 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

166

167 Ala Arg Asn Thr Thr His Tyr Arg Ala Asn Trp Val Arg Gly Pro Gly
168 -190 -185 -180 -175

169

170 Ser Trp Thr Ser Pro Pro Leu Leu Leu Ser Thr Cys Pro Gly Ala
171 -170 -165 -160

172

173 Ala Arg Ala Pro Arg Leu Leu Tyr Ser Arg Ala Ala Glu Pro Leu Val
174 -155 -150 -145

175

176 Gly Gln Arg Trp Glu Ala Phe Asp Val Ala Asp Ala Met Arg Arg His
177 -140 -135 -130

178

179 Arg Arg Glu Pro Arg Pro Pro Arg Ala Phe Cys Leu Leu Arg Ala
180 -125 -120 -115

181

182 Val Ala Gly Pro Val Pro Ser Pro Leu Ala Leu Arg Arg Leu Gly Phe
183 -110 -105 -100 -95

184

185 Gly Trp Pro Gly Gly Gly Gly Ser Ala Ala Glu Glu Arg Ala Val Leu
186 -90 -85 -80

187

188 Val Val Ser Ser Arg Thr Gln Arg Lys Glu Ser Leu Phe Arg Glu Ile
189 -75 -70 -65

190

191 Arg Ala Gln Ala Arg Ala Leu Gly Ala Ala Leu Ala Ser Glu Pro Leu
192 -60 -55 -50

193

194 Pro Asp Pro Gly Thr Gly Thr Ala Ser Pro Arg Ala Val Ile Gly Gly
195 -45 -40 -35

196

197 Arg Arg Arg Arg Arg Thr Ala Leu Ala Gly Thr Arg Thr Ala Gln Gly
198 -30 -25 -20 -15

199

200 Ser Gly Gly Gly Ala Gly Arg Gly His Gly Arg Arg Gly Arg Ser Arg
201 -10 -5 1

202

203 Cys Ser Arg Lys Pro Leu His Val Asp Phe Lys Glu Leu Gly Trp Asp
204 5 10 15

205

RAW SEQUENCE LISTING
PATENT APPLICATION US/09/945,182DATE: 01/11/2002
TIME: 22:33:10

INPUT SET: S36718.raw

206 Asp Trp Ile Ile Ala Pro Leu Asp Tyr Glu Ala Tyr His Cys Glu Gly
207 20 25 30
208
209 Leu Cys Asp Phe Pro Leu Arg Ser His Leu Glu Pro Thr Asn His Ala
210 35 40 45 50
211
212 Ile Ile Gln Thr Leu Leu Asn Ser Met Ala Pro Asp Ala Ala Pro Ala
213 55 60 65
214
215 Ser Cys Cys Val Pro Ala Arg Leu Ser Pro Ile Ser Ile Leu Tyr Ile
216 70 75 80
217
218 Asp Ala Ala Asn Asn Val Val Tyr Lys Gln Tyr Glu Asp Met Val Val
219 85 90 95
220
221 Glu Ala Cys Gly Cys Arg
222 100
223

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1207 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: Homo sapiens

(vii) IMMEDIATE SOURCE:

- (B) CLONE: MP52

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 845..1204

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

246
247 ACCGGGCGGC CCTGAACCCA AGCCAGGACA CCCTCCCCAA ACAAGGCAGG CTACAGCCCG 60
248
249 GACTGTGACC CCAAAGGAC AGCTTCCCGG AGGCAAGGCA CCCCCAAAAG CAGGATCTGT 120
250
251 CCCAGCTCC TTCCTGCTGA AGAAGGCCAG GGAGCCCGGG CCCCCACGAG AGCCCAAGGA 180
252
253 GCCGTTTCGC CCACCCCCCA TCACACCCA CGAGTACATG CTCTCGCTGT ACAGGACGCT 240
254
255 GTCCGATGCT GACAGAAAGG GAGGCAACAG CAGCGTGAAG TTGGAGGCTG GCCTGGCCAA 300
256
257 CACCATCACC AGCTTTATTG ACAAAGGGCA AGATGACCGA GGTCCCGTGG TCAGGAAGCA 360
258

PAGE: 1

SEQUENCE VERIFICATION REPORT
PATENT APPLICATION US/09/945,182

DATE: 01/11/2002
TIME: 22:33:11

INPUT SET: S36718.raw

Line	Error	Original Text
------	-------	---------------

PAGE: 1

SEQUENCE MISSING ITEM REPORT
PATENT APPLICATION US/09/945,182

DATE: 01/11/2002
TIME: 22:33:11

INPUT SET: S36718.raw

<< THERE ARE NO ITEMS MISSING >>

PAGE: 1

SEQUENCE CORRECTION REPORT
PATENT APPLICATION US/09/945,182

DATE: 01/11/2002
TIME: 22:33:11

INPUT SET: S36718.raw

Line	Original Text	Corrected Text
1945	(C) INDIVIDUAL ISOLATE: primer number 8	(C) INDIVIDUAL ISOLATE: primer number 8